

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 7, 2001, 00:18:23 ; Search time 54.1 Seconds  
(without alignments)  
212.181 Million cell updates/sec

Title: US-09-494-297-2  
Perfect score: 3945  
Sequence: 1 MKKTRFPNKLTNTQTVLS.....IAGISLGWIGHTIRIKHD 757

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 297327 seqs, 15163773 residues  
Total number of hits satisfying chosen parameters: 297327

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents, AA, New:  
1: /cgnl\_7/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgnl\_7/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgnl\_7/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgnl\_7/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgnl\_7/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgnl\_7/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3945	100.0	757	5	US-09-494-297-2
2	1958.5	49.6	742	5	US-09-494-297-4
3	138.5	3.5	625	5	US-09-692-064-2
4	127	3.2	795	5	US-09-430-590E-142
5	121.5	3.1	1027	4	US-08-956-171C-5254
6	121.5	3.1	2329	5	US-09-421-124-16
7	121.5	3.1	2329	5	US-09-421-124-16
8	120	3.0	1312	1	PCT-US01-08117-95
9	120	3.0	3418	5	US-09-421-124-44
10	120	3.0	3418	5	US-09-421-124-44
11	119.5	3.0	886	4	US-08-956-171C-5235
12	118	3.0	708	5	US-09-792-024-123
13	117	3.0	995	5	US-09-430-590E-14
14	117	3.0	5373	1	PCT-US01-03782A-96
15	117	3.0	5447	1	PCT-US01-03782A-284
16	115	2.9	1992	4	US-08-945-567C-3
17	113.5	2.9	1833	4	US-08-945-567C-4
18	112.5	2.9	903	5	US-09-193-562D-46
19	111.5	2.8	1090	1	PCT-US01-04098A-3733
20	110	2.8	1068	1	PCT-US01-04098A-1890
21	109.5	2.8	1021	1	PCT-US01-04098A-1765
22	109	2.8	520	5	US-09-792-024-121
23	109	2.8	846	5	US-09-591-178-33
24	108.5	2.8	774	5	US-09-739-449-13266
25	108	2.7	1561	5	US-09-456-474-23
26	107	2.7	1308	1	PCT-US01-03782A-383
27	107	2.7	1494	5	US-09-421-124-186

28	107	2.7	1494	5	US-09-421-124-186	Sequence 186, App
29	107	2.7	1501	1	PCT-US01-03782A-312	Sequence 312, App
30	105	2.7	803	1	PCT-US01-04098A-987	Sequence 987, App
31	105	2.7	931	1	PCT-US01-04098A-1311	Sequence 1311, App
32	105	2.7	949	1	PCT-US01-04098A-3279	Sequence 3279, App
33	105	2.7	1576	5	US-09-430-590E-5	Sequence 5, App1
34	104.5	2.6	542	5	US-09-270-849B-194495	Sequence 194495,
35	104.5	2.6	1579	5	US-09-421-124-184	Sequence 184, App
36	104.5	2.6	1579	5	US-09-421-124-184	Sequence 184, App
37	104.5	2.6	2568	6	US-60-248-823-86	Sequence 86, App1
38	104	2.6	801	1	PCT-US01-11797-29	Sequence 29, App1
39	104	2.6	1088	1	PCT-US01-04098A-3858	Sequence 3858, App
40	103.5	2.6	751	1	PCT-US01-04098A-1607	Sequence 1607, App
41	103.5	2.6	802	1	PCT-US01-04098A-3575	Sequence 3575, App
42	103.5	2.6	1423	5	US-09-196-296B-2	Sequence 2, App1
43	103	2.6	1008	5	US-09-308-453-2	Sequence 2, App1
44	103	2.6	1180	5	US-09-559-497-42	Sequence 42, App1
45	101.5	2.6	1339	1	PCT-US01-04098A-1083	Sequence 1083, App

## ALIGNMENTS

RESULT 1  
US-09-494-297-2  
; Sequence 2, Application US/09494297  
; GENERAL INFORMATION:  
; APPLICANT: POBBLESKI, ANDREAS  
; TITLE OF INVENTION: COLLAGEN-BINDING PROTEINS FROM STREPTOCOCCUS PYOGENES  
; FILE REFERENCE: P06628050/AS  
; CURRENT APPLICATION NUMBER: US/09/494, 297  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 757  
; TYPE: PRT  
; ORGANISM: Streptococcus pyogenes  
US-09-494-297-2

*Copy*

Query Match	100.0%	Score 3945	DB 5	Length 757
Best Local Similarity	100.0%	Pred. No. 2.7e+290		
Matches 757	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MKKTRFPNKLTNTQTVLSKSKRFTVTLVGYFLMIFALVTSVNGAKYFGLVESSTPN	60	
DB	1	MKKTRFPNKLTNTQTVLSKSKRFTVTLVGYFLMIFALVTSVNGAKYFGLVESSTPN	60	
QY	61	AINDPDSSEYKRWYGYEYVNGHRYKOFRAHDLRVNLBESRSYQYCFNLKAFPLGSD	130	
DB	61	AINDPDSSEYKRWYGYEYVNGHRYKOFRAHDLRVNLBESRSYQYCFNLKAFPLGSD	130	
QY	121	SSYKWKYKKHKGISTFKEDYAMSPRITGDELNOKLRVYVNGHPOANGIMEGLEPNAI	180	
DB	121	SSYKWKYKKHKGISTFKEDYAMSPRITGDELNOKLRVYVNGHPOANGIMEGLEPNAI	180	
QY	181	RVQEAIVWYSDNAPISNPDESFRSESNLVSTQSLMRQALKQIDPDLTKMPKQV	240	
DB	181	RVQEAIVWYSDNAPISNPDESFRSESNLVSTQSLMRQALKQIDPDLTKMPKQV	240	
QY	241	PDPFQSLTFSESEKQKYNKGYONLSGGLVPTKPPPGPPMPNPQPTTVLIRKAI	300	
DB	241	PDPFQSLTFSESEKQKYNKGYONLSGGLVPTKPPPGPPMPNPQPTTVLIRKAI	300	
QY	301	GDYSKLEGGATLDTGDNVNSFOARVSSNDIGERIELSGYTLRELNSPAGYSTAEP	360	
DB	301	GDYSKLEGGATLDTGDNVNSFOARVSSNDIGERIELSGYTLRELNSPAGYSTAEP	360	
QY	361	TFVEVAGKYVTIIDGKQIENPKIEVPEYSEVAYNDFEESVLTTONYAKFYAKKNGS	420	
DB	361	TFVEVAGKYVTIIDGKQIENPKIEVPEYSEVAYNDFEESVLTTONYAKFYAKKNGS	420	

```

OY      421  SQVYVCFNADLKSPPDSDDGKMTMPDTTGEVKYTHIAGRLPFYTVKPRDTDDTFLK  480
Db      421  SQVYVCFNADLKSPPDSDDGKMTMPDTTGEVKYTHIAGRLPFYTVKPRDTDDTFLK  480
OY      481  HIKRVIEKGYREKGOALIEYSGLTETQLRAATOLAIYPTDSAEIDKRLKDYHGFGDND  540
Db      481  HIKRVIEKGYREKGOALIEYSGLTETQLRAATOLAIYPTDSAEIDKRLKDYHGFGDND  540
OY      541  STLAVAKILVEYAQDSNPOLDLDFIIPNNKYSLLGTQMHPPDLDVDIRMEKKEVI  600
Db      541  STLAVAKILVEYAQDSNPOLDLDFIIPNNKYSLLGTQMHPPDLDVDIRMEKKEVI  600
OY      601  PVTNHLTRKIVTGLAGRTKDFHEIIELNKKOELLQTVKDKTNLEFKDKRATINLK  660
Db      601  PVTNHLTRKIVTGLAGRTKDFHEIIELNKKOELLQTVKDKTNLEFKDKRATINLK  660
OY      661  HGESLTGLLEGYSYLVKEPDSSEGYKRVKVSQEVANAATVSKTGITSDETLAFENKKEPV  720
Db      661  HGESLTGLLEGYSYLVKEPDSSEGYKRVKVSQEVANAATVSKTGITSDETLAFENKKEPV  720
OY      721  VPTGYDOKINGYIALIYIAGISLGIWGIHTIRIKHD  757
Db      721  VPTGYDOKINGYIALIYIAGISLGIWGIHTIRIKHD  757

RESULT      2
US-09-494-297-4
: Sequence 4, Application US/09494297
: GENERAL INFORMATION:
: APPLICANT: POBIELESKI, ANDREAS
: TITLE OF INVENTION: COLLAGEN-BINDING PROTEINS FROM STREPTOCOCCUS PYOGENES
: FILE REFERENCE: P06628U0/BAS
: CURRENT APPLICATION NUMBER: US/09/494,297
: CURRENT FILING DATE: 2000-01-31
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 742
: TYPE: PRT
: ORGANISM: Streptococcus pyogenes
: US-09-494-297-4

```

[illegible]

Db	358	QVENPNKEVAPRYEVEAYNDFMDEEVLGCTFPYGFXYATNKKRKSQVYVCFNADLHSP	417
Qy	436	DSEGGCTMPDPFTT-GEVKYTHAGADLTKRYVKKPNDTPDPFFLKIIRKIVIEKGYREKG	494
Db	418	DSYSGELINPDUSTMKEVVKTHAGSDLEKRYALRPDPJNPEDFLKHIIKRIEKGYYKKK	477
Qy	495	QAIYESGLTEQLAQAOLAIYPTDSAELEDKDL---KDYGFQGMNSTLAVAKILV	550
Db	478	DS--YNGLTIEQFPAALQOLAIYTTDSADLTKLKTYNNGKGYYGFESMDKTLAVIKELI	535
Qy	551	EYAADSNPQTLDDPFELPNNKKYQSLLGTQWPEDELVDIIIRMEKK-EVAPYTHMLTLR	609
Db	536	TYAANGSAPQTLNDDPFYPPNNSKQSLIGIECHGDDLDVDIRMEKKQOEVIPTVTHSLTVK	595
Qy	610	KTVYGLAGDRTKDFEHFEIELKNNKOELLISQTVKTDKTNLEPKDQKATINLKHBSLTLQG	669
Db	596	KTVVDELDAKTKGFQFELELTKDKTCGPVIMTLKTNMDDLAKQDQKYSFNLKHBDHTIRIEG	655
Qy	670	LPEGYSLVLEKTDSEGGKRYKVAQNSQEVANAIYVSKTGLTSDETLAFENKKEVYTVQYDKI	729
Db	656	LPTGYSTYLKAEAKKDYIIVYDNKRVSQEASVGDITIEDKRVTFENKRKDLVPPGLTDDG	715
Qy	746	NGYLAIVYAGISLGIW	
Db	716	AIYMLILVPLGLIYW	732

```

RESULT 3
US-09-692-064-2
; Sequence 2, Application US/09692064
; GENERAL INFORMATION:
; APPLICANT: Minion, F. Chris
; APPLICANT: Menon, Sreekumar A.
; APPLICANT: Mahairas, Gregory G.
; TITLE OF INVENTION: VACCINE ADJUVANT
; FILE REFERENCE: 08411-016001
; CURRENT APPLICATION NUMBER: US/09/692,064
; PRIORITY FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US 60/160,429
; PRIORITY FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 625
; TYPE: PRF
; ORGANISM: Mycoplasma hyopneumoniae
US-09-692-064-2

```

[illegible]

```
QY 288 PQTTSYRYYAAGDYSKLEGCATLOLTGDNVNSFOARFSSNDGEIETLSDGYTTLTE 347
      : | | | : | | : | | |
Db 259 -----FIKIKEYLLTYAKI-----ETDFINEPE-----283
      : | | | : | | : | | |
QY 348 LNSPAGYSIAEPIITFEVGAQVYTIIDGKOENPKKEIVEPYSEAYAMDEEFVSVALTON 407
      : | | | : | | : | | |
Db 284 -----KINKIRISAEIAN-KVNYIDVYDKSIWMDSD-----313
      : | | | : | | : | | |
QY 408 YAKFYAAKN-----KNGSSQVYVCFNADLKSPDSED-GGKTPTPPPTTGVEKYT 456
      : | | | : | | : | | |
Db 314 --KNMAKKNDHPHSIOGYRKTAHOLLKLDOEEKDSDSAELKNTFTNFDPDFENKPYS 371
      : | | | : | | : | | |
QY 457 HIAGRDLFFYTVPRDTPDTFLKHIKIVIEKREKGCALESGLTFEQLRATOLAIY 516
      : | | | : | | : | | |
Db 372 KV-----IDLVSFAKSKEPLEK-LMENNOTSSEFINAKST-----405
      : | | | : | | : | | |
QY 517 YFTDSAEJLDKDLKDYHGCGDMNDSLAVAKILIVEYAODSNAPOLDLDF--FIPNNKKY 574
      : | | | : | | : | | |
Db 406 FDDDOEAIRAD--DKRTGENT-----VREIV-----SLPIPDNDEFRELLPVKNPF 449
      : | | | : | | : | | |
QY 575 QSLIGTOWMHED---LVQIIRMEDKK-----EVLPVTNHLTRKVYVGLAGDRTKD 622
      : | | | : | | : | | |
Db 450 VKALINSYLCKPAGSLINXIOLEENKVDYARPNIKIFDTIIDSTRKMAVFA-----E 504
      : | | | : | | : | | |
QY 623 PHFEIELKNNKQ--ELLISOTVKT----DKTNLEFKDG--KATINIKHGESLTLGOPL 671
      : | | | : | | : | | |
Db 505 LMNDOEIKERKMSPOLLFTLRNALISPDLTKL--KDSATFKIIMLNKPCPOIITLGLS 562
      : | | | : | | : | | |
QY 672 EGYSYLUKE---TDSEGKRVKNSOEVANAATSKGITISDETLEFENNKCEVPVPT 723
      : | | | : | | : | | |
Db 563 KPTSVKPEKPKDOSSKPODOTJISSQSÖSGTSS-----TDSKTATTENÖKPAEDOT 612
```

```

RESULT      4
US-09-430-590E-142
; Sequence 142, Application US/09430590E
; GENERAL INFORMATION:
; APPLICANT: Poultier, et al.
; TITLE OF INVENTION: UNSUAL RETROTRANSPOSON FROM THE YEAST CANDIDA ALBICANS
; FILE REFERENCE: 674521-2001.1
; CURRENT APPLICATION NUMBER: US/09/430,590E
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,342
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 142
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: sequence of retrotransposon from unknown organism
US-09-430-590E-142

Query Match      3.2%; Score 127; DB 5; Length 795;
Best Local Similarity 19.2%; Pred. No. 0.43;
Matches 127; Conservative 103; Mismatches 238; Indels 194; Gaps 377

OY      75 YESVGRCHPYKQFRAVHDLRVNLDEGSRGYOVCENLKK---APPLGSDSSVKKWKKKH 130
      |  :: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      223 YAAHNRHNPRIKEF-----DGATPYERK-YGLSKYVYIPFQFTDVLV-KCASVQ 270

OY      131 DGISTKR---EDYAMSPRITG-----DELNQKLRAYVYNGHPONANGIMEGLEPLNAI 180
      |  :: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      271 EALSIKLPSRDKAFTVWGAFLGSGDSFTFRV-LVSTKGIVYIT---SNIRPIATM 326

OY      181 RVTQEAWVYSDNAPISNPDESEFKRESNELVSTQSLIMRQALKQLIDPMLATKMPKQV 240
      |  :: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      327 QVLNDVYLAIVSENSISY--DDTF-----LSPLNHPMIRTN----- 360

OY      241 PDGFQLSFESEKCGKYNKGIONLISGLVPTKKPPTPED-PPMPRPNOPTTSVLIRKYA 299
      |  :: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Db 361 -----QHDRGDNMINVEYEN-----RPNVFHYHAEPRNTSGTIDRP-- 400
QY 300 IGDYSKLLEGATLDLTGDNVNSFOARFYSSNDIGERIELSDGYTL----- 345
Db 401 --DIRPADPTWQMRPDANIHQ--ETTVQTPDRHGLDTMINNEHQLPERSGEGNPGQOVR 457
QY 346 -----TELNSPAGYSIAEPTTFVEYAGKYVTLIDGKOIENPNKEIPEYSVEA 393
Db 458 TDIIGFRDRGPTLPLNTPIDGVDP-----ETDDI--SMSENIDSPNEMIIISPLT 510
QY 394 YNDE-----EFSVLTQO-----NYAKFYAAKKNGSSOYVYCFNADLKSPD 436
Db 511 -NELHEHIDISSGMSLQTNMEADNELKTNEMVLYKSKN-DGIIIOOQOFTENTL--SDEN 567
QY 437 SEDGKTWTPDFTTGGEVYTHIAGRDLEFKYTVKPRD-----TDPRFLKHIRKVIK 488
Db 568 EED---SSTDELTLEDKQOURL-----EYNISNDMINNDQONEDTQVPHAKENY 618
QY 489 GYRKGGQA---IF---YSGLTETQLRAATQALAYVYTTDSAEILDKOKLKYDHGFGGND 541
Db 619 ETQSRNETNMRIMGIIENLSDGKNTPRELRIVYTDNKKREIK-----YQDS 667
QY 542 TLAAKTLIVEAODSNPOLDLDFIPRNKKXOSLGTOMHPED--LVDIRM---EDK 597
Db 668 NIEL-----SEPRNENENOTFIESN--LELLNQEMFOEDRPOVEDIRLTTRKKK 715
QY 598 EVIP---VTNHLTLRKTYVTLAGDRDQHFHEIELKNNKOKELLSQLVKT--DK--TNLEFKD 652
Db 716 SLSPDFNQTHN-----ETQLPMADINEDMLEEYEDENINMEVLADSTETLDKELDEES 770
QY 653 GK 654
Db 771 GR 772

```

RESULT 5  
 US-08-956-171C-5254  
 Sequence 5254, Application US/08956171C  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 Gil H. Choi  
 Patrick S. Dillon  
 Steven A. Rosen  
 Steven C. Barash  
 Michael R. Fannon  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/956,171C  
 FILING DATE: 20-Oct-1997  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/009,861  
 FILING DATE: January 5, 1996  
 APPLICATION NUMBER: 08/781,986  
 FILING DATE: January 3, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hoover, Kenley K.  
 REGISTRATION NUMBER: 40,302  
 REFERENCE/DOCKET NUMBER: PB248P1  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 610-5790  
 TELEFAX: (301) 309-8439  
 INFORMATION FOR SEQ ID NO: 5254:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1027 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5254:  
 US-08-956-171C-5254

Query Match 3.1%; Score 121.5; DB 4; Length 1027;  
 Best Local Similarity 20.2%; Pred. No. 1.7; Indels 225; Gaps 27;  
 Matches 133; Conservative 72; Mismatches 230;

153 OKLRAYVNGHPONANGIME--GLEPLNAIRVTOEAVWYSDNAPISNDESFKRES 209  
 49 OKTTVEENGSATDNKTSSTOTATNVNHEETQ-----SYNATVT-----EDPSNA 96  
 210 NLVSTSLIMRALKQIDPNLATKMPKQVDDFQSLFSEEDKGDYKQYQNLISG 269  
 97 TOVTEAPRAVQAPQAPANIEYKEEVKEAKRQVKEETQSD--NSGDQROYD-- 152  
 270 LVPTKPTPDPMPPNOPTSVLRKYAIGDYSKLLEGATLQLTGDNVNSFOARFSS 329  
 153 LTPKK-----ATONQVAETQVEVAQPTASEK-----PRVTR 186  
 330 NDIGERIELSDGYTLTLENSPAGYSIAEPTFEKVEAGKYTTIDGKQIENPKETVEPY 389  
 187 ADVAEAKESNA-----KVEGTGTDVT-----SKVTEIG 215  
 390 SVRAYNDEEFSVLTQNYAKFYAKRKNSSQVYVCFMNDL-----KSPDS 437  
 216 SIEGHNTNVEPHAGRAVLKYLKFNGLHOGDY-FDETLNNVNTGVSARAKYPEI 274  
 438 EDGKMTPTDFTTGEVYTHIAGRLFKYVPRDPDFLKHKKVIE-KGYREKQA 496  
 275 KNSGVV---ATGEVLE---GKIRITFTNDIEDKYDTAELEINFLDPKVVQNGNQ 327  
 497 IEVSGLETLQRAATOLAIYFTDSALD--KDKLDYHFGSDMNSTLAVAKILVEYA 553  
 328 TITSTINEEO-----TSKELDVYKDGIGNY--VANLMS-----IET 364  
 554 QDSNPQLDLDFFIRNNKYOSLI-----GPO-----WHPREDVDII- 591  
 365 NKAN-NRFSVAFIKPNNGKTSVTVGTLMKGSNONGNPKVRIEYLGNNEDIKSVY 423  
 592 -RMEDKKEVLPVTHNLIRKTYGLAG-----DRTKDFHEIE----- 628  
 424 ANTTDTSKFEKVSNNMGNLONNGYSYININLTKTYVNHGDELNGTIDEVPTQM 483  
 629 -----LKNKKOELLQTVKTDKTNL 648  
 484 VGHPEOLYKYRGTLLTDNGLVLYSKNANGNEKNKPIQNNKEFKYKDTIKETLQG 543  
 649 EFDGKATILKKGESITLQ-----GLPEGSYLVKRETDSCG---TKVYNSQ 693  
 544 YDKNLVTVVEEY-DSSTLIDIDYHTALDGGGYVDGYETIEETDSALDIDYHTAVDSE 602

RESULT 6  
 US-09-421-124-16  
 ; Sequence 16, Application US/09421124  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Futreal, Phillip A  
 ; APPLICANT: Wooster, Richard F  
 ; APPLICANT: Ashworth, Alan  
 ; APPLICANT: Stratton, Michael R  
 ; TITLE OF INVENTION: Materials and methods relating to the  
 ; TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer  
 ; TITLE OF INVENTION: susceptibility gene and uses thereof.

NUMBER OF SEQUENCES: 222  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Bell Seltzer Park & Gibson  
 STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107  
 CITY: Raleigh  
 STATE: NC  
 COUNTRY: USA  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/421,124  
 FILING DATE:

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/755,587  
 FILING DATE: 25-NOV-1996  
 APPLICATION NUMBER: GB 9523959.6  
 FILING DATE: 23-NOV-1995  
 APPLICATION DATA:  
 APPLICATION NUMBER: GB 9525555.0  
 FILING DATE: 14-DEC-1995  
 APPLICATION DATA:  
 APPLICATION NUMBER: GB 9617961.9  
 FILING DATE: 28-AUG-1996  
 APPLICATION DATA:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kenneth D Sibley  
 REGISTRATION NUMBER: 31,665  
 REFERENCE/DOCKET NUMBER: 5405-135  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2329 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-421-124-16

Query Match 3.1%; Score 121.5; DB 5; Length 2329;  
 Best Local Similarity 17.9%; Pred. No. 6.7; Indels 305; Gaps 35;  
 Matches 138; Conservative 104; Mismatches 224;

133 ISTKFEDYAMSPRITGD-----ELNOKL----- 155  
 504 ISRGKESYKMSDKLKNNGVESDVELTKNIPMEKNQVCALENKYKVELLPREKYRVAS 563  
 156 --RAVWNGHPONANGIMEGLEPLNAIRVTOEAVWYSDNAPIS---NPD-ESFRRESE 208  
 564 PSRKVOFN--QNTN-----LRVIOKNQE-----ETTSISKITVNDSEELFSDBE 606  
 209 SNLY-----STQSLSIMRQA 223  
 607 NNFVFOVANNERNNLALNTEKELHETDLTCVNEPIFNKSTWVLYGTGDKQATVDSIKKDL 666  
 224 LKQIDPN-----LATKMPKQVDDFQSLFSEEDKGDYKQYQNLISGVLPTKRP 276  
 667 VYVLAENKNSVQKHIMLIGDOLKSDISLNT---DKIPEKNNDNMNMAAGLGP----- 718  
 277 TPEDPAPNPQOTSVLIRKYAIGDYSKLLEGATLQLTGDNVNSFOARVSSNDIGERI 336  
 719 -----ISNHSFGGSFRASNKETIKLSHNK--KSKMF-FKDIIEQY 757  
 337 ELSDGYTLTLENSPAGYSIAEPTFEKVEAGKYTTIIDGKQIENPKETVEPYVEAYND 396  
 758 -----PTSLACVE-----IVNTLALDN--OKLSKPOSITVSA 789  
 397 FEEFSVLTQNYAKFYAKRKNSSQVYV---FNADLSPDS-----EDGK 442  
 790 HLOSVVSD-----CKNSHTLPQMLFSKODFNHNHNLTPQKQITELSTLSDSGS 842  
 443 TWPDPFTTGEVYTHIAGRLFKYVPRDPDFLKHKKVIEK-----G 489

```

Db 843 QF--EFTQFR-KPSYLLQKSTFEV-----PENQMTILKTSEBCROADLHVIMNAPS 891
QY 490 YREKGAIEYSGLTETOLRAATOL-----AIYFTDSALDKDKLYHGFQDMNDST 542
Db 892 IGVDSKQFEGTVEIKRFRAGILKNDCKNSASGYLTDENEVGFGRFYSAHG-TKLNVST 950
QY 543 LAVAKTLVEYAQDSNPQOLDLDFIPNNKYSLSLGTQMHPRDLYDIIMEKKEVIVP 602
Db 951 EALQKAVKLFSDIENISEETSAEV-----HPISLSSKCH-DSVVSMKIEIEN----- 996
QY 603 THNLTRKTYTGLAGDRTKDFHEIELKNNKOBLLSQT-----VTKDTNLE 649
Db 997 -HN-----DKT-----VSEKNNKCOLLIQNNIEMTTGTFVEIEITENYKRNT 1037
QY 650 FKDGKATINLKHGESLTLOGLPBGYSYLV--KE-----TDSBGYKVKVNSQ----- 693
Db 1038 NEDNKYTAASRNSHNLFEFSDSSKNDYVCIHKDEDTLFTDOHNICLKLSGQFMKEGNT 1097
QY 694 -----EVANATVSKTGITSDTELAFENKKEPVVPTGVQDKINGX 732
Db 1098 QIKEDLSDLTFLEVAKAQEACHGNTS-----NKEQLATATEQNIKDF 1140

```

## RESULT 7

US-09-421-124-16

Sequence 16, Application US/09421124

GENERAL INFORMATION:

APPLICANT: Futrell, Phillip A

APPLICANT: Wooster, Richard F

APPLICANT: Ashworth, Alan

APPLICANT: Stratton, Michael R

TITLE OF INVENTION: Materials and methods relating to the

TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer

NUMBER OF SEQUENCES: 222

CORRESPONDENCE ADDRESS:

ADDRESS: Bell Seltzer Park &amp; Gibson

STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107

CITY: Raleigh

STATE: NC

COUNTRY: USA

ZIP: NC 27622

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,124

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/755,587

FILING DATE: 25-NOV-1996

APPLICATION NUMBER: GB 9523959.6

FILING DATE: 23-NOV-1995

APPLICATION NUMBER: GB 9525555.0

FILING DATE: 14-DEC-1995

APPLICATION NUMBER: GB 9617961.9

FILING DATE: 28-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kenneth D Sibley

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5405-135

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 2329 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-421-124-16

Query Match 3.1%; Score 121.5; DB 5; Length 2329;  
 Best Local Similarity 17.9%; Pred. No. 6.7;  
 Matches 138; Conservative 104; Mismatches 224; Indels 305; Gaps 35;

```

QY 133 ISTKFEYAMSPRITGD-----ELNOKL----- 155
Db 504 ISRGKESYKMSDKLKGNNESVDYELTKNIPMEKNODYCALNENYKNVLLPPEKYARVAS 563
QY 156 --RAVYNGHPQANGIMELEPLNARIYQEAWVWYSDNAPIS---NPD-ESFKRESE 208
Db 564 PSRKQGFN--QNTN-----LVIQKNQ-----ETTSISKITVANDSELSDSNE 606
QY 209 SNLY-----STQSLMRA 223
Db 607 NNFFQVQANERNMIALGNTKELHETDLCVNEBIFPNSTWLYGDTGDKATQVVSIRKDL 666
QY 224 LKQIDPN-----LATKPKQVPPDDFOLSTPESDEKDGKYGKGYONLLSGGLVPTKPP 276
Db 667 VYVLAENKNSVQHAKMTIGQDKSDISLNT---DKLPEKNNIDYNNKAGLGP---- 718
QY 277 TPGDPPMPNPQOTTSLYLRKVAIGDYSKLLEGATLQTLGDNVNSFOARVSSNDIGERI 336
Db 719 -----ISNHSFGSGFRFASNKELKSEHNK--KSKMF-EKIDEGY 757
QY 337 ELSDGTYTTLTSLNPAQYSAEPIITRVEAGKYTTIIDGQIENPKIEIPEYSVEAYND 386
Db 758 -----PTSLACVE-----IVNTLALN--OKIKSPQISINVSA 789
QY 397 FEEFSVLTQNVAKFYAKKNSOYVC---FNADLKSPDS-----EDGK 442
Db 790 HLOSSVYVSD-----CKNSHTPOMLFESKOPNSHNHNTPSQKEQITLSTILDSGS 842
QY 443 TMTPTDTEGKYYTHIAGRDLFKYTKPRDTPDFELKHAKYIEK-----G 489
Db 843 QF--EFTQFR-KPSYLLQKSTFEV-----PENQMTILKTSEBCROADLHVIMNAPS 891
QY 490 YREKGAIEYSGLTETOLRAATOL-----AIYFTDSALDKDKLYHGFQDMNDST 542
Db 892 IGVDSKQFEGTVEIKRFRAGILKNDCKNSASGYLTDENEVGFGRFYSAHG-TKLNVST 950
QY 543 LAVAKTLVEYAQDSNPQOLDLDFIPNNKYSLSLGTQMHPRDLYDIIMEKKEVIVP 602
Db 951 EALQKAVKLFSDIENISEETSAEV-----HPISLSSKCH-DSVVSMKIEIEN----- 996
QY 603 THNLTRKTYTGLAGDRTKDFHEIELKNNKOBLLSQT-----VTKDTNLE 649
Db 997 -HN-----DKT-----VSEKNNKCOLLIQNNIEMTTGTFVEIEITENYKRNT 1037
QY 650 FKDGKATINLKHGESLTLOGLPBGYSYLV--KE-----TDSBGYKVKVNSQ----- 693
Db 1038 NEDNKYTAASRNSHNLFEFSDSSKNDYVCIHKDEDTLFTDOHNICLKLSGQFMKEGNT 1097
QY 694 -----EVANATVSKTGITSDTELAFENKKEPVVPTGVQDKINGX 732
Db 1098 QIKEDLSDLTFLEVAKAQEACHGNTS-----NKEQLATATEQNIKDF 1140

```

## RESULT 8

PCT-US01-08117-95

Sequence 95, Application PC/TUS0108117

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: YUE, Henry

APPLICANT: LAL, Preeti

APPLICANT: LU, Dzung Anna M.

APPLICANT: PATTERSON, Chandra

APPLICANT: AZIMZAI, Yalda

APPLICANT: BANDMAN, Olga

APPLICANT: TANG, Y. Tom

APPLICANT: MATTHUR, Preete



```

QY 345 LTELNSPAGYSIAEPIFKYEAGK-----VYTIIDGKOINPKKEIPEYSVE 392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1942 VSK1-SECDVSLSTDCICKSIGIKLHKSVSASANTCGIFSTASGSKSVSPASLONAROV- 1999
QY 393 AYNDFFESVLTTONAKFYAKNNKSSQVYCFNADLKSPD--SEDC-----GKTM 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2000 -----FSEIEDSTKQVFSKVLFSKNEH-SDQLTRENTAIRTEPHILISQKFSYVNVVSSA 2054
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 445 TPDFTTGEVRYTHIAGDLFKYTVKPRDPTDFLKHIKVIEKREKQALIEYSGLTE 504
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2055 FSGFSTASGKQVSLSSSLHKV-----KGLVEEFDLIR----- 2087
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 505 TOLRAATQALIIYFTDS-----AELDKDKLDY-----HGF 535
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2088 -----TEHSLHYSPTSROVSKILPRVDRKRNPEHCYNSMEKTSKFKLSNNLNEGG 2141
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 GDMNDSTLAVAKLIVEAODSNPPQLDLPFIPINNKQYSLIGTQHPDLDVLIIMED 595
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2142 SSENNSIKVSPYLSQFOO-----KQDLVLGK-----VSLVENIHVLG 2181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 KKEVIPTVHNLTKRYT-----GLADRTKDF--HFETELKNNKQELLISQTVKT 643
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2182 KEQASPKVMEIGKTEFSDVPVKTNIIEVCSYKSDSENYFEFEAVEIKAFAEDELDT 2241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 644 DK-----TNLEFKDKKATINLKHGESLTLQGLPESYSLVYKETDSEGYKVK 689
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2242 DSKLPSHATSLFTCPBENEMVLSNSRIGKRGPPLVLGEPISIKRNLLNEFD---RII 2297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 690 VNSOEVANATVSKTGITSDETL 711
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2298 ENOEKSLKASKS-----TPDGTI 2315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

# RESULT 10 US-09-421-124-44

```

: Sequence 44, Application US/09421124
: GENERAL INFORMATION:
: APPLICANT: Futreal, Phillip A
: APPLICANT: Wooster, Richard F
: APPLICANT: Asmworth, Alan
: TITLE OF INVENTION: Materials and methods relating to the
: TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
: TITLE OF INVENTION: susceptibility gene and uses thereof.
: NUMBER OF SEQUENCES: 222
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Bell Seltzer Park & Gibson
: STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
: CITY: Raleigh
: STATE: NC
: COUNTRY: USA
: ZIP: NC 27622
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,124
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/755,587
: FILING DATE: 25-NOV-1996
: APPLICATION NUMBER: GB 9523959.6
: FILING DATE: 23-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9525555.0
: FILING DATE: 14-DEC-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9617961.9
: FILING DATE: 28-AUG-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Kenneth D Sibley

```

```

: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5405-135
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3418 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-421-124-44

```

Query Match 3.0%; Score 120; DB 5; Length 3418;  
Best Local Similarity 18.3%; Pred. No. 16;  
Matches 125; Conservative 102; Mismatches 259; Indels 196; Gaps 27;

```

QY 119 SDSSVKKWKRRKHDISTKFDYAMSPRITGDELNOKLRAVYNGHPQ---NANGIMEGLE 175
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1741 SNSSMSNSYTHSD-----EYNDNRILSKNKLDSGIEPLVKNEDQKNFSFKVISNVK 1795
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 PLNAIRVT-----QBAWYISDNAPISNPDESFKRE-SESNLVSTLSQSLMRQALKOL 227
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1796 DANAYPQTVNEDICVEELV---TSSSPCKNNKNAIKLISNSNFEVGPAPAFRIASGKIR 1852
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 IDPMLATKMPKQVDDPDLQSLFSESEDKDKYKNGYONLISGLVPTKPPPGDPMPMPNQ 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1853 LCHSHETIKKAYDI-DSRSKIYFTKENNENK-SKICQTKIACGE----- 1895
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 PQTTSVLIRKAYIGDYSKLEGATLQLTGDNVNSFOARVS---SNDIGERIELSDGTYT 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1896 -----ALDSEDLIHN---LDNDECSHSHKVFADIQSEIILQHNQMSGLEK 1941
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 345 LTELNSPAGYSIAEPIFKYEAGK-----VYTIIDGKOINPKKEIPEYSVE 392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1942 VSK1-SECDVSLSTDCICKSIGIKLHKSVSASANTCGIFSTASGSKSVSPASLONAROV- 1999
QY 393 AYNDFFESVLTTONAKFYAKNNKSSQVYCFNADLKSPD--SEDC-----GKTM 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2000 -----FSEIEDSTKQVFSKVLFSKNEH-SDQLTRENTAIRTEPHILISQKFSYVNVVSSA 2054
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 445 TPDFTTGEVRYTHIAGDLFKYTVKPRDPTDFLKHIKVIEKREKQALIEYSGLTE 504
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2055 FSGFSTASGKQVSLSSSLHKV-----KGLVEEFDLIR----- 2087
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 505 TOLRAATQALIIYFTDS-----AELDKDKLDY-----HGF 535
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2088 -----TEHSLHYSPTSROVSKILPRVDRKRNPEHCYNSMEKTSKFKLSNNLNEGG 2141
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 GDMNDSTLAVAKLIVEAODSNPPQLDLPFIPINNKQYSLIGTQHPDLDVLIIMED 595
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2142 SSENNSIKVSPYLSQFOO-----KQDLVLGK-----VSLVENIHVLG 2181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 KKEVIPTVHNLTKRYT-----GLADRTKDF--HFETELKNNKQELLISQTVKT 643
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2182 KEQASPKVMEIGKTEFSDVPVKTNIIEVCSYKSDSENYFEFEAVEIKAFAEDELDT 2241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 644 DK-----TNLEFKDKKATINLKHGESLTLQGLPESYSLVYKETDSEGYKVK 689
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2242 DSKLPSHATSLFTCPBENEMVLSNSRIGKRGPPLVLGEPISIKRNLLNEFD---RII 2297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 690 VNSOEVANATVSKTGITSDETL 711
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2298 ENOEKSLKASKS-----TPDGTI 2315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 11 US-08-956-171C-5235

```

: Sequence 5235, Application US/08956171C
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: APPLICANT: Gil H. Choi
: APPLICANT: Patrick S. Dillon
: APPLICANT: Craig A. Rosen
: APPLICANT: Steven C. Barash

```

Michael R. Fannon  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171C  
 FILING DATE: 20-Oct-1997  
 CLASSIFICATION: <unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861  
 FILING DATE: January 5, 1996  
 APPLICATION NUMBER: 08/781,986  
 FILING DATE: January 3, 1997

## ATTORNEY/AGENT INFORMATION:

NAME: Hoover, Kenley K.  
 REGISTRATION NUMBER: 40,302

## REFERENCE/DOCKET NUMBER: PB248P1

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 610-5790  
 TELEFAX: (301) 309-8439

## INFORMATION FOR SEQ ID NO: 5235:

## SEQUENCE CHARACTERISTICS:

LENGTH: 886 amino acids  
 TYPE: amino acid

STRANDEDNESS: single  
 TOPOLOGY: linear

## MOLECULE TYPE: protein

## SEQUENCE DESCRIPTION: SEQ ID NO: 5235:

US-08-956-171C-5235

Query Match 3.0%; Score 119.5; DB 4; Length 886;  
 Best Local Similarity 19.3%; Pred. No. 1.9;  
 Matches 139; Conservative 100; Mismatches 259; Indels 221; Gaps 36;

QY 92 HDL-----RVN-----LEGSRYQYVCNFKKAFPLGSDSSVKKWKRRD-----GIST-----K 136  
 DB 118 HDIGPREQVNFOLLDKNNETQYHF-----FSIKPADVYYTRKKAVELDINTASTWKK 172  
 QY 137 FEDYAMSPRITGDELNOKL--RAVMYNGHPON-----ANGIMEGLEPLNATRVQ-----E 185  
 DB 173 FEYV-----ENNOKLPLRVLSYSPVEDIAYTRFPVSDOTOLKITYSSQIDGEE 223  
 QY 186 AVWYSD--NAPISNPDESFKRESSENLVSTQSLMRQALKQLIDPILATMPKQVPD 242  
 DB 224 TNYDYTKLVFAKPIYNDPSLVKSDTNDAYVTDQSS----- 259  
 QY 243 DQQLSFSEEDKDKYKNGYQNLISGLVP---TKRPEDPMPMPNPQOTTSVLIRKXA 299  
 DB 260 ----SVASNOTNTNTSNTSTINNANNOPOATTNMSOPAPSSSTNADASSQPAHE-- 313  
 QY 300 IGDYKSLBGAATLQTDGVNVSFOARVSSNDIGERIELSDGTYYTLTELNSPAGYIAEP 359  
 DB 314 ----TNSGNTNDKTNESSNQ-----SD-----VNOQPPPADDESQDA 347  
 QY 360 IFEKVEAGVYVITIDKQIENPKKEIVEPYSVEAYNDFEESVLTQNTAKFYANKNG 419  
 DB 348 I-----KNPATIDKEHPADNWRPI-----DFOKNNDKGEHGFHYASTVE 387  
 QY 420 SSOVYVCFNADLKSPPDSEGGGKTMP---DFTTGEVKTTHLAGNDLFFYTKKPRDTP 475  
 DB 388 PATVIF-----TKTGPIELGLKTASTWKKFEVYEGDKLP---VELVSY-----DSDK 433

QY 476 DTFKHKIKVIEKGYREGQALIEYSGLETQLRATQALITFTDSAELEDKDKYHGF 535  
 DB 434 D--YAVIFPVSNGTRE-----VKIVSISYGENIRHBDY--- 465  
 QY 536 GDMNSTLAVAKILVEYAO--DSNPQLDLDDEFIPNNKKYOSLGTOMHPEDLYD--II 591  
 DB 466 ---DYTLNV-----FAPITNPDDYVDEET-----NLQKL-APYHKAKTIERQVY 509  
 QY 592 RMEDKREVIPTVHNLTRKVTYGLAGDRTPDHFIELEKNNKOELLQOTVTKDTNLEPK 651  
 DB 510 ELEKIQEKIPEKTKAEYKKKL-----DQTR-----VELADQKSAVTEBENTPTNDOLF 559  
 QY 653 DCK-----ATINLKHEBSLTGLPE-----GTYLVKTEDSGKYVKYNSQEVANA 698  
 DB 560 DLQEAHFVYFESEENSESYMDGFVEHPFYATLNGQKYVMATKDSYWKDLIVGKRYT 619  
 QY 699 TVSKTGISDETLAFENKNEPYPVPGVDKINGYALAIYAGISGIMGHTIRIRKHD 757  
 DB 620 TVSKDPKNSRTLIF-----PIIP--DKAVYNAIVKVVANI--GYEQYHVRILIND 668

## RESULT 12

US-09-792-024-123

Sequence 123, Application US/09792024

## GENERAL INFORMATION:

APPLICANT: Roemer, Terry

APPLICANT: Jiang, Bo

APPLICANT: Boone, Charles

APPLICANT: Bussey, Howard

TITLE OF INVENTION: Gene Disruption Methodologies for Drug

FILE REFERENCE: 10182-004-999

CURRENT APPLICATION NUMBER: US/09/792,024

CURRENT FILING DATE: 2001-02-20

NUMBER OF SEQ ID NOS: 490

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 123

LENGTH: 708

TYPE: PRF

ORGANISM: Candida albicans

US-09-792-024-123

Query Match 3.0%; Score 118; DB 5; Length 708;  
 Best Local Similarity 18.6%; Pred. No. 1.7;  
 Matches 128; Conservative 108; Mismatches 251; Indels 202; Gaps 31;

QY 136 KREDYAMSPRITGDELNOKLRVMYNGHPONANGIMEGLEPLNATRVQEAWYVS---D 192  
 DB 68 KIEDRGILF--IDDDGLKQVLRD---NSKSSNDTRRAHRTDTLVDITDGNOMHTSHQ 121  
 QY 193 NAPISNPDESFKRESSENLVSTQSLMRQALKQLIDPILATMPKQVPD---DFQLSI 248  
 DB 122 DKQISFEDKMEHEERD-----VPIINODYFKVYSPN---NQPSIDDKTRKQGDYIF 172  
 QY 249 FESEDKGDKYKNGYQNLISGLVPTKPPGDPMPMPNPQOTTSVLIRKXAIGDYSKLE 308  
 DB 173 KNNDDKOKAERCGK----- 189  
 QY 309 GATLQLTGD---NVNSFOARVSSNDIGERIE--LSDGTYYTLTELN-----SPAGYS 355  
 DB 190 SIYAEVLVKNLPASIESEFNNTYLLSDRLSNENFORKSLISALSFEKGTDSITGHE 249  
 QY 356 IAPPIFEKVEAGVYVITIDKQIENPKKEIVEPYSVEAYNDFEESVLTQNTAKFYANKNG 412  
 DB 250 ISLIKNNLGHGQKF--LIFSLSKNANDE---YTLEDETHIELNLSQAFKSGGLFYCP 304  
 QY 413 -----YAKNKNSSQ-----VYVCFNADLKSPPDSEGGGKTMP---TPDF----- 448  
 DB 305 GMFLVIEGITSASGANSNDHGIGGCFYVSNIGHPSEERSRSTLDYVGNLDELGNHROI 364  
 QY 449 --TTGE-----VAKYTH---IAGNDLFRYTKKPRDTPDFTLKIHKVY 485

DB 365 APTGSKIRIKSKKKRLVLEKTLNKLIFVGTDLVLDKRLVLDGRLKFKQKLENSI 424  
QY 486 IEKGYREKCO-----AIEYSGLETQLRAATQALAIYFTDSAE-----523  
DB 425 IESIEDEECOMAGCTNIPALVPTGTSFVSFPLSVTNSVTNITNSYSKSNFDNFTIYS 484  
QY 524 -----LDKRLADYHGFQGMNDSTLAVAKILVYADSNPQ-----LDDDFITNNKQYS 576  
DB 485 KVPNTYSRCKIILIPKKNPMOSTYSLGSSSLMYFPQSSIPKFIKRLKRLKGN-----540  
QY 577 LIGTQHP-----EDLV-----DII---RMEDKEVIPTVHNLTKRTVT 613  
DB 541 -LVSNMPTNINLSQELVYFDELTAKLRNDIIFPRIOQEBELIADODORTNEERIN 599  
QY 614 GLAGDRTKDFHEFIELKNNKQELLSQVKT--DKTNLE-FKDGKATINKHGESLTQGL 670  
DB 600 NLT--QNKNTHLPSKTKQAR-----KLVTILDQNLQPLKMLKLINLAYDYSLEIPL 652  
QY 671 PEGSYLVETDSEGYKVKVNSQEVANAT 699  
DB 653 PS--VIIINDSSFDNFEVYNGCKVYVNI 679

RESULT 13  
US-09-430-590E-14  
Sequence 14, Application US/09430590E

GENERAL INFORMATION:  
APPLICANT: Poulet, et al.  
TITLE OF INVENTION: UNUSUAL RETROTRANSPON FROM THE YEAST CANDIDA ALBICANS  
FILE REFERENCE: 674521-2001.1  
CURRENT APPLICATION NUMBER: US/09/430,590E  
CURRENT FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/106,342  
PRIOR FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 156  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 14  
LENGTH: 995  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: sequence of retrotransposon from unknown organism  
NAME/KEY: UNSURE  
LOCATION: (1)...(995)  
OTHER INFORMATION: 'XAA' can be any amino acid  
US-09-430-590E-14

Query Match 3.0%; Score 117; DB 5; Length 995;  
Best Local Similarity 18.2%; Pred. No. 3.6;  
Matches 123; Conservative 103; Mismatches 238; Indels 210; Gaps 37;

QY 75 YESYVGHVYKQFRVRAHDLRLVLEGSRSYQVCFNLK-----APFLGSSSVKKYKKH 130  
DB 418 YAVHINHPRIKE-----DGAIPERY-YGLSKYVIPFGQYDVLV--KCAVQ 465  
QY 131 DGISTKFE-----EDYAMSPTIG-----DELNQLRAVYNGHPONANGIMEGLEPLNAT 180  
DB 466 EALSLPLSRDCAFTVYFGAFLGYSDSFTPRV-LVSTKGTPTVTT--SMIRIATM 521  
QY 181 RYQEAENVYISDNAPISNPDESFKRESSENLVSTQSLSLMAKQKOLIDNLTATKMPKV 240  
DB 522 QVLDVLAIVSESSISY-DDTF-----LSPLNHPMIRTN-----555  
QY 241 PDDFQSLFSESEDKGKYNKGYONLSGLVPTKPPTPGD-PRMPNQGQTSVLLRKYA 239  
DB 556 -----QHDRGNINVEYN-----KPNVPFEYIAEPPRTNNSSTGIIDRP--555  
QY 300 IGDYSLLECATQLQTDVNVNSFOARVFSNDIGERIELSDGTVTL-----345  
DB 596 --DIRRADPTWQMRPDANIHQ-ETTVQTPRDELDPTMINNHQLRPRSEGVNPGQVRA 652

QY 346 -----TELNSPAGYSIAEPTTFKVEAGKYVTIIDQKQENPKIEVPEYSVEA 393  
DB 653 TDIIGQFRDRGPTTLNTPIDLGVPD-----ETDDI-SMSENPIDSPSEMITSPLP 705  
QY 394 YNDFE-----EFSVLTQO-----NYAKFYAKNKSQVYCFNADLKSPD 436  
DB 706 -NELEHQIIDSSEMSILQINMEADNELKTNEMVILKSKN-DGIITIIQQQOFTENL--SDEN 762  
QY 437 SEDGKTMTPTDFTTEGKYTHIAGRDLFKYTVKPRD-----TDDPFLKIKKIVIEK 488  
DB 763 EED-----SSTDEETLEDKQOORL-----EYNISPNDWINNDVQNEDDIQVPHVK-----808  
QY 489 GYREKQATEYSLTETQLRAATQ-----AIYFTDSA-----EL-----DKKIKD 531  
DB 809 -----EPIN-----ETQSHNGTNPRIENGITENLSDOGKNTPRELRVYTYDNKKIKQ 858  
QY 532 Y-----HGGDMNDSTLAVAKILVEYADSNPQDLDLDFIPNNKQY 575  
DB 859 YKNSNIEISFRBNKNHTIEN-----LELDDQEMQEDPOVEDIRLTTPKKDKSL 912  
QY 576 SLIGTQHPEDLVDIRMEDKKEVI-PVTHNLTKRTVTGLAGDRTKDFHEIELKNNKQ 634  
DB 913 SPDFNQTHNE--IQLPWADINEDMLEYDENINMNEVLAD--STETLDRKELDDESSGRI 968  
QY 635 ELLSQTVKTDKTNL 648  
DB 969 EYIADRVX-KTEV 981

RESULT 14  
PCT-US01-03782A-96  
Sequence 96, Application PC/TUS0103782A

GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-040  
CURRENT APPLICATION NUMBER: PCT/US01/03782A  
CURRENT FILING DATE: 2001-02-05  
PRIOR APPLICATION NUMBER: 09/496,914  
PRIOR FILING DATE: 2000-02-03  
PRIOR APPLICATION NUMBER: 09/598,075  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: 09/620,325  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 60/250,583  
PRIOR FILING DATE: 2000-11-30  
NUMBER OF SEQ ID NOS: 386  
SOFTWARE: Custom  
SEQ ID NO 96  
LENGTH: 5373  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-03782A-96

Query Match 3.0%; Score 117; DB 1; Length 5373;  
Best Local Similarity 19.9%; Pred. No. 59;  
Matches 145; Conservative 121; Mismatches 258; Indels 206; Gaps 36;

QY 84 YKQFRVADLRVNLG--SRSTQVYCFNLKAFPLGSSSVKKYKKHDKGISTKREDYA 141  
DB 2323 YKTLGVTLHERQESLQAIINRMEVH--KEA-----NSVQWQESKEV-LKSDAM 2371  
QY 142 MSPRITGD-----ELNQLRAVYNGHPONANGIMEGLEPLNATIRTOEAENVYISDNAP 196  
DB 2372 SSTKTEYVKAQAESKAPLAEL-----EONSPIQVKEKALGLVTV-----PN 2417  
QY 197 SNPDESFKR-----ESESNT-----VSTSQ-----SLMR 221  
DB 2418 SQEAENMKRIQELNSRWERATEVTVARQRLBESASHLACFOAASQSLAPMLKEKELM 2477  
QY 222 QALQOL-IDPNLTATKMPKQVPPDFQSLIFES-EDKGDYIKKGYQNLISGLVPTTKRPTG 279

```

Db 2478 GVLGLSIDPNNLNAKQOV--QFMKEFEARQOHEOLNEAQCILTG-----PG 2526
Qy 280 DPMPPNPQPTTSVLRKXAYIDYSKLEGAFLQLTGDNVNSFOARVSSNDIGERIELS 339
Db 2527 DVSLSSTQVOKELQISINO-----KWVELT-DKLNRSISOI--DOAIVSTQYQ 2571
Qy 340 DGTYYLTTELNSPAG--YSTAEPTFEKVENAGKYTTIIDGKOINPNK-----EIVEPYSVE 392
Db 2572 ELLQDLSEKVRVAVGRLSVQSAISTQPEAVK-----QOLETSEIRSDLLQDLHEVKE 2624
Qy 393 AYNDPEEPSVLTTQNYAKFYAKN-----KNGSSQVYVCFNADLKSPDSE--- 438
Db 2625 AQTLCDELISVLIGEOLYKDELKRLKETVALPLQGLEDLAADRINRLQALASTQOFOQMF 2684
Qy 439 DQGTMTPTDFTTGEVKYTHIAGRDLEKTYVPRDNDPDTFLKHIKKVIEKGYREGQALIE 498
Db 2685 DELRTWLDKQSOQAKNCPISSAK-----LERLQSOLOENEFQKSLNQ 2727
Qy 499 YSGLTETQLRAATQALAIYFTTSABELDKDK-----LKDYHGFQDNSTLAVAKIL 549
Db 2728 HSGSYEVIVAAEBESLIL-----SVPRGEEKRTLQNLQVLELKNH--WEELSKTTADRQSR 2780
Qy 550 VEYADDSNPQLTDLDFIIPNNKYOSLIGTQWHPEDVDII-----RMEDKKEVI-PVT 603
Db 2781 KDCMOKA-----OKY-----QWVEDLVPMIEDCKAKMSELRTLDPVQ 2819
Qy 604 HNLTLRKYTVGLAGDRTKDFHEIELKNNKOBELLSQTVTKDNTLEFKDGKATIN----- 658
Db 2820 LSSSLRSKAML--NEVERRSLETLINSADLILNSSEDEDCI--RDEKAGINONMDA 2875
Qy 659 -----LKHG--ESLT--LQGLPEGISYLKETSDESGYKVVY-----NSQEVANATVSKTG 704
Db 2876 VYEELOAKTGSLEEMTORLREFQESFKNEKVEGAKHQLEIFDALGSQACSNKNLEKLR 2935
Qy 705 ITSDETLAFE 714
Db 2936 AOEVLQALE 2945

RESULT 15
PCT-US01-03782A-284
; Sequence 284, Application PC/TUS0103782A
; GENERAL INFORMATION:
; APPLICANT: Byseq, Inc.
; APPLICANT: Ford, John E et al
; TITLE OF INVENTION: Novel Bone Marrow Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-040
; CURRENT APPLICATION NUMBER: PCT/US01/03782A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/250,583
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: Custom
; SEQ ID NO 284
; LENGTH: 5447
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US01-03782A-284

```

Query Match 3.0%; Score 117; DB 1; Length 5447;

Best Local Similarity 19.9%; Pred. No. 60;

Matches 145; Conservative 121; Mismatches 258; Indels 206; Gaps 36;

Qy 84 YXKQFVAHDLRVNLEGG--SRSYQVYCFNLKKAAPLGSSSVKRWYKKHGDISTKFEDEYA 141

```

Db 2396 YEKLGVLHEROESLOAILNREEVH---KEA-----NSVLQWLESKEEV-LKSMAM 2444
Qy 142 MSPRITGD-----ELNOKLRVAVNGHPONANGIMEGLEPLNAIRVTOEAVWYSDNAPI 196
Db 2445 SSPRTETVYKQAESKNKAFLEL-----EONSPKIQKVKEALAGLLVY-----PN 2490
Qy 197 SNPDSEFKR-----ESESUL-----VSTSQL-----SLMR 221
Db 2491 SOEAENMKIOELNSRWERATEVTVAROLOEESASHLACFOAESQLOQPMLEKELM 2550
Qy 222 QALKQL-IDPNLATKMPKQVPPDDFQLSIFES-EDKGDYKNKGQYMLSGGLVPTKPPPG 279
Db 2551 GVLGPLSIDPNNLNAKQOV--QFMKEFEARQOHEOLNEAQCILTG-----PG 2599
Qy 280 DPMPPNPQPTTSVLRKXAYIDYSKLEGAFLQLTGDNVNSFOARVSSNDIGERIELS 339
Db 2600 DVSLSSTQVOKELQISINO-----KWVELT-DKLNRSISOI--DOAIVSTQYQ 2644
Qy 340 DGTYYLTTELNSPAG--YSTAEPTFEKVENAGKYTTIIDGKOINPNK-----EIVEPYSVE 392
Db 2645 ELLQDLSEKVRVAVGRLSVQSAISTQPEAVK-----QOLETSEIRSDLLQDLHEVKE 2697
Qy 393 AYNDPEEPSVLTTQNYAKFYAKN-----KNGSSQVYVCFNADLKSPDSE--- 438
Db 2698 AQTLCDELISVLIGEOLYKDELKRLKETVALPLQGLEDLAADRINRLQALASTQOFOQMF 2757
Qy 439 DQGTMTPTDFTTGEVKYTHIAGRDLEKTYVPRDNDPDTFLKHIKKVIEKGYREGQALIE 498
Db 2758 DELRTWLDKQSOQAKNCPISSAK-----LERLQSOLOENEFQKSLNQ 2800
Qy 499 YSGLTETQLRAATQALAIYFTTSABELDKDK-----LKDYHGFQDNSTLAVAKIL 549
Db 2801 HSGSYEVIVAAEBESLIL-----SVPRGEEKRTLQNLQVLELKNH--WEELSKTTADRQSR 2853
Qy 550 VEYADDSNPQLTDLDFIIPNNKYOSLIGTQWHPEDVDII-----RMEDKKEVI-PVT 603
Db 2854 KDCMOKA-----OKY-----QWVEDLVPMIEDCKAKMSELRTLDPVQ 2892
Qy 604 HNLTLRKYTVGLAGDRTKDFHEIELKNNKOBELLSQTVTKDNTLEFKDGKATIN----- 658
Db 2893 LSSSLRSKAML--NEVERRSLETLINSADLILNSSEDEDCI--RDEKAGINONMDA 2948
Qy 659 -----LKHG--ESLT--LQGLPEGISYLKETSDESGYKVVY-----NSQEVANATVSKTG 704
Db 2949 VYEELOAKTGSLEEMTORLREFQESFKNEKVEGAKHQLEIFDALGSQACSNKNLEKLR 3008
Qy 705 ITSDETLAFE 714
Db 3009 AOEVLQALE 3018

```

Search completed: June 7, 2001, 00:22:29  
Job time: 246 sec